DOCUMENT RESUME

ED 050 146 TM 000 533

AUTHOR	Hedl, John J., Jr.; And Others
TITLE	Computer-Based Intelligence Testing.
INSTITUTION	Florida State Univ., Tallahassee. Computer-Assisted Instruction Center.
PUB DATE	Feb 71
NOTE	22p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Attitude Tests, Automation, College Students, *Computer Programs, Evaluation, *Intelligence Tests, Item Analysis, Psychological Testing, Scoring, *Test Construction, *Testing, Test Scoring Machines, *Test Validity
IDENTIFIERS	SIT, *Slosson Intelligence Test, WAIS, Wechsler Adult Intelligence Scale

ABSTRACT

The development and evaluation of an interactive computer-based administration and scoring program for the Slosson Intelligence Test (SIT) is discussed. A computer program successfully developed for administering and scoring the Wechsler Adult Intelligence Scale is cited to support the feasibility of developing this proposed SIT program. Testing, scoring, content analysis procedures, and validity evaluation of the computer-based approach, are described. Student input to the IBM 1500 Instructional System, used for the administration of the computer-based SIT (CB-SIT), consisted of typewritten responses which were automatically recorded and evaluated. Anxiety scales were administered by the computer system before and after the CB-SIT, and attitude scales were presented in the conventional paper-and-pencil form. Validation of the CB-SIT was conducted by a comparative evaluation in which the SIT and WAIS were traditionally administered. A Latin square design was used to counterbalance potential order and sequence effects of the three test administrations. Data analysis involved evaluation of the test administration procedure and determination of concurrent validity. Results indicated high correlation between the CB-SIT and the original form and confirmed that an automated administrative system is both feasible and valid for SIT. (LR)

ED050148

COMPUTER-BASED INTELLIGENCE TESTING

John J. Hedl, Jr., Harold F. O'Neil, Jr., and Duncan N. Hansen

The Florida State University

Paper read at the American Educational Research Association Convention, New York, 1971

COMPUTER-BASED INTELLIGENCE TESTING[1]

John J. Hedl, Jr., Harold F. O'Neil, Jr., and Duncan N. Hansen
Florida State University

The purpose of this research was to develop and evaluate an interactive computer-based administration and scoring program for the Slosson Intelligence Test (Slosson, 1963). Although individual intelligence testing remains a major time-consuming endeavor for school and clinical psychologists, no interactive computer programs have been developed to administer, score, or interpret intelligence test data. Given that the assessment of intellectual functioning dominates a major portion of a psychologist's time commitment, and that intelligence test scores are still our best predictor of school success (Vernon, 1958), a reliable and valid computer-based approach has immediate applications for psycho-educational assessment services.

The feasibility of such a technological approach has been demonstrated by recent findings in this area of evaluation. For example, Elwood (1969) has developed an automated testing booth programmed to administer the Wechsler Adult Intelligence Scale (Wechsler, 1955). The testing apparatus is capable of automatically administering all verbal and performance subtests of the Wechsler Adult Intelligence Scale (WAIS). The system records the patient's responses to the individual test items, scores the Digit Symbol and Digit Span subtests, and provides several temporal measures for item responses.

Griffith and Elwood (1969) have reported favorable test-retest reliability coefficients for the automated WAIS administration (Verbal IQ (VIQ), $r = .99$, Performance IQ (PIQ), $r = .97$, and Full Scale IQ (FS), $r = .99$). Orr (1969) compared the automated WAIS with a traditional administration of the WAIS. Concurrent validity coefficients between the two administration methods were as follows: For VIQ's, $r = .94$; for PIQ's, $r = .82$. The correlation between FS-IQ's was .93. The main problem concerning the comparability of the two methods seems to lie primarily in the performance section.

The initial WAIS data are quite encouraging and demonstrate the feasibility of automated individual intelligence test administrations. However, this non-computerized system only provides scoring capabilities for two of the eleven subtests (Digit Span and Digit Symbol), and does not attempt to evaluate student verbal responses on a real-time basis. The significant feature of our approach to automated individualized intelligence assessment is that the entire diagnostic evaluation occurs within a terminal-oriented interaction between the student and the computer system. In this respect, a more optimum simulation of the interaction found within the traditional psycho-educational testing situation may be achieved. We shall now turn to a description of our program developmental activities.

Description of Developmental Activities for the Computer-Based Slosson Intelligence Test

The Computer-based Slosson Intelligence Test (CB-SIT) has been designed to operate with an IBM 1500 Instructional System (IBM, 1967). Terminals of this system consist of a cathode ray tube (CRT), a light pen, and a typewriter keyboard. All student responses and response latencies are automatically recorded by the CAI system. The test items have been developed from the ages of 12 to 27. This range appears adequate for an assessment of a college student's intelligence (IQ).

The student reads the item presented on a CRT and responds by typing in his response for immediate computer evaluation. If the student's answer is judged as totally correct or incorrect, the program continues to the next item. If the student is partially correct, he is instructed by the computer to more fully explain his response. The presentation strategy essentially matches the recommended method of standardized administration. The development of the scoring program was different for the two main types of items of the SIT. For the items requiring math answers, the author devised the answer sets. The answer sets were then coded with standard CAI techniques. Those items requiring verbal responses, however, required other techniques which are reported below.

A major problem for an automated natural language program is that of synonymity (the expression of equivalences among words and sentences). A two-phase strategy was employed to empirically construct the answer analysis algorithms.

First, the vocabulary portion of the Slosson Intelligence Test (SIT) was administered via paper and pencil techniques to 84 college students enrolled in an introductory habilitative science course. These obtained responses were then scored by a skilled psychologist according to scoring criteria outlined in the SIT test manual. Second, the correct answers were content analyzed and categorized according to similarities. In essence, the first phase defined the fundamental attributes that would determine a correct word definition. The second phase focused on the manner in which these concepts could be alternately expressed.

A specific example will clarify this procedure. The test word "plutocracy" contains two basic concepts in the standard meaning, that of government or control, and that pertaining to the rich or wealthy class of society. By using the sample data, and by searching Roget's International Thesaurus (1962) and Soule's Dictionary of English Synonyms (1966), two lists of word entries were constructed. Each dictionary set of entries contains those words which can be used to correctly describe each of the two concepts for the test item. The student, in order to obtain a correct answer, must simultaneously match a word in each of the two concept categories. However, if the student matches an entry in one category, he is given another opportunity to answer the item. This procedure for keyword analysis largely ignores the syntactic relationships of the verbal input.

For other words, the procedure for analysis was similar. In certain instances, particular phrases are stored in the analysis buffers which must be matched to obtain a correct answer. This procedure of categorizing the test items into concepts or components has enabled efficient program modifications to be implemented where necessary.

It should be obvious that the present answer sets do not contain all possible lexical elements for describing an answer. A heuristic principle was employed in the design of the answer dictionaires. For example, 90 percent of the sample answers to the test item "environment" were composed of the word "surroundings." This, then, defines the first analysis check point for the program. The most prevalently observed answers are analyzed first by the program before other subroutines are used. In this respect, the developmental procedures focused on program efficiency based on observed frequencies of responses.

For the first major field trial of the program, 25 students (14 males and 11 females) were obtained from an introductory psychology class at FSU during the Spring Quarter, 1970. As part of the developmental procedures, all students responded to the entire set of test items from the ages of 12 to 27. In order to investigate the psychological impact of this interactive testing situation, the A-State scale of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, and Lushene, 1970) was used to measure levels of state anxiety before and after the test situation. In addition, an attitudinal measure was given concerning the automated testing session.

Student test item protocols obtained from detailed computer performance records were typed and given to a skilled psychologist for scoring of each item. A psychologist derived IQ scores from his individual evaluation of each student's performance data. This scoring procedure will be referred to as the psychologist-derived method of evaluation. The data was then compared to the computer-derived method of evaluation which consisted of another psychologist deriving IQ scores based on the scoring program's decisions per item. From the computer listing of correct or incorrect answers, the psychologist derived IQ scores in the manner specified in the SIT test manual.

A major question concerning the adequacy of the programmed answer analysis procedures can be answered by viewing the power of these scoring procedures for detecting correctly the adequacy of the input responses. A computer may score an answer as correct or incorrect and may in either case be correct or incorrect from a clinician's viewpoint. The decision outcomes, in other words, may be a valid positive (the joint occurrence of correct classifications of an answer by the computer and the psychologist), a false negative (where the computer classifies an answer as incorrect whereas in reality it is correct as judged by a psychologist, a valid negative (where both the computer and the psychologist score the answer as incorrect), or a false positive (where the computer scores an answer as correct when the psychologist scores the answer as incorrect).

In a comparison of the computer-scored answers, the SIT scoring program correctly scored 87 percent of the responses as either valid positives or valid negatives. Of the 215 errors committed for the entire 1681 test responses, 207 or 95 percent were false negatives. Only 5 percent of the errors were false positives. For items requiring word or sentence input, the correct response classification rate (hit rate) was 85, whereas the hit rate was 91 percent for items requiring numerical input. Program modifications were implemented based upon these data. In [illegible] changes consisted of enlarging the item dictionaries [illegible] rect answers.

The correlation between the computer-derived IQ scores and the psychologist-derived IQ scores was .83. The mean IQ for the psychologist-derived IQ scores was 129 (sd=8.7). In comparison, the mean of the computer-derived IQ scores was 118 (sd=8.2). The small standard deviations were probably due to the homogeneous nature of our college student population.

No significant elevation in state anxiety was found when an analysis of the before and after anxiety data was performed. Moreover, student comments obtained during a debriefing session indicated that this interactive testing procedure seemed to have led to favorable attitudes. Thus, the computer testing experience seemed to be perceived as nonthreatening.

In summary, the results strongly indicate the feasibility of an interactive computer-based method for the assessment of intelligence. However, the validity of this approach must still be demonstrated.

Validation of Computer-Based Slosson Intelligence Test

The second phase of our developmental effort consisted of a validity evaluation of this computer-based approach to individualized intelligence testing. Specifically, the validity evaluation attempted to determine: (1) the relationship between a computer-based administration of the Slosson Intelligence Test (SIT) and a traditional administration of the SIT, and (2) the concurrent validity of these two SIT test scores in relationship with the Wechsler Adult Intelligence Scale (WAIS).

In order to more explore the complex nature of this man-machine testing application, student attitudes and anxiety were assessed both before and after the different testing situations.[2]

Method

Subjects

Forty-eight undergraduate students were obtained from courses at Florida State University. S sources included the introductory psychology and educational psychology classes. Ss were randomly assigned to the three experimental order conditions with an equal number of males and females within each group. Participation in the investigation served to fulfill certain course requirements for the students.

Apparatus

The computer-based SIT was administered via an IBM 1500 Instructional System (IBM, 1967). Terminals for this system consist of a cathode ray tube (CRT), a light pen, and a keyboard. These terminals are located in an air-conditioned, sound-deadened room. The CAI system also administered the anxiety scales before and after the CB-SIT. The attitude scales were presented in conventional paper and pencil fashion. All student input for the CB-SIT consisted solely of typewritten responses. All student responses were automatically recorded and evaluated by the CAI system.

Standard testing materials were employed for the WAIS and SIT administrations.

Experimental Design

A Latin square design (Winer, 1962) was employed to counterbalance potential order and sequence effects of the three different intelligence test administrations. Each S was individually tested with the WAIS, SIT, and CB-SIT. A time period of approximately one week separated the different test administrations. In addition, each testing session included an administration of an attitude scale and anxiety scale both before and after each testing period. Figure 1 exemplifies the design of the investigation.

	Week 1	Week 2	Week 3
Group 1 (N=16)	WAIS	CB-SIT	SIT
Group 2 (N=16)	CB-SIT	SIT	WAIS
Group 3 (N=16)	SIT	WAIS	CB-SIT

Fig.1.--Experimental design with number of Ss

Ss were randomly assigned to the three experimental testing groups with an equal number of males and females within each group. In order to control for possible examiner effects, there was a random assignment of WAIS and SIT examiners to students.

All Ss reported to the FSU CAI Center for all three experimental testing sessions. The CB-SIT was presented via an IBM 1500 Instructional System. Ss were asked to read a description of the operation of the CAI terminal. They were then given practice in the operation of the terminal keyboard and instructed how to carry out erase and enter functions. After "signing on," a number of practice frames were presented to familiarize the Ss with the terminal characteristics and operations.

The administrations of the WAIS and SIT conformed carefully to the standard procedures as described in their respective test manuals. The testing rooms were free from distracting noises and intrusions. These rooms were well-lighted and ventilated. Sufficient time was scheduled for the testing so that good rapport could be established and maintained and that the administration could proceed in an easy, unhurried manner.

The WAIS examiners were five graduate students from the Department of Psychology at the Florida State University. The four SIT examiners were graduate students from the Department of Educational Research and Testing at Florida State University. All examiners have had extensive experience with their respective intelligence tests.

Upon completion of the experimental procedures for all tests, the Ss were debriefed, given some additional information concerning the general nature of the experiment, and cautioned not to discuss the experiment with other Ss who had not yet completed the procedures.

Results

The data analyses for the present investigation consisted of: 1) an evaluation of the test administration procedure, and 2) correlational techniques to determine the concurrent validity between the WAIS, SIT, and CB-SIT.

Evaluation of Test Administration Procedures

The means and standard deviations for the WAIS FS, SIT, and CB-SIT IQ scores for this sample of college students (N=48) are presented in Table 1. As may be seen in Table 1, the mean WAIS FS-IQ scores are fairly consistent across all three experimental testing orders. However, CB-SIT and SIT IQ scores do not show the same consistent relationship, i.e., in Groups I and II there is a consistent increase from the two administrations of the SIT, whereas in Group III, there is a decline in SIT scores. Therefore, a 3 x 3 Latin Square repeated measures analysis of variance (Winer,

TABLE 1.--Means and Standard Deviations of WAIS, SIT, and CB-SIT Test Scores by Entire Sample and for Experimental Group

GROUP	TESTS		
Group I	WAIS	CB-SIT	SIT
$\bar{X}$	119.7	123.7	135.4
sd	5.9	8.6	9.4
Group II	CB-SIT	SIT	WAIS
$\bar{X}$	119.5	132.4	121.3
sd	8.4	9.3	6.0
Group III	SIT	WAIS	CB-SIT
$\bar{X}$	127.1	121.0	120.2
sd	9.1	8.1	8.7
Entire Sample	WAIS	CB-SIT	SIT
$\bar{X}$	120.7	121.2	131.6
sd	6.7	8.6	9.7

1962) was employed to evaluate test, order, and sequence effects. The dependent variable was the obtained IQ scores for the WAIS, SIT, and CB-SIT-derived scores.

The results of this analysis revealed a statistically significant difference between the IQ scores obtained from the WAIS, SIT, and CB-SIT (F=74.9, 2 and 90 *df*, $p < .01$). In addition, a significant order effect of test administration was found (F=8.38, 2 and 90 *df*, $p < .01$). Sequence of administration effects and test x order interaction were not found to be statistically significant.

Concurrent Validity Evaluation

Pearson product-moment coefficients were used to determine the validity relationships between the WAIS, SIT, and CB-SIT (computer-derived scores). Since a significant effect of test order was found, the concurrent validity data will be presented for both the entire sample, as well as for three experimental testing orders, i.e., Groups I, II, and III.

Table 2 presents the correlation coefficients between the WAIS, SIT, and CB-SIT for the entire sample of college students. These results

TABLE 2.--Coefficients of Correlation Between WAIS, SIT, and CB-SIT (N=48)

TEST	CB-SIT	SIT	WAIS VERBAL IQ	WAIS PERF. IQ	WAIS FS-IQ
CB-SIT		.75**	.54**	.32*	.54**
SIT			.48**	.32*	.52**

* $p < .05$

** $p < .01$

indicate that the CB-SIT and SIT correlate positively with each other ($p < .01$). In addition, both the CB-SIT and SIT correlate positively with WAIS VIQ's ($p < .01$), WAIS PIQ's ($p < .05$), and WAIS FS-IQ's ($p < .01$). The magnitudes of the correlation differences between the correlations of the CB-SIT and SIT with the WAIS were tested for statistical significance using a series of Fisher t tests for correlated samples (McNemar, 1962).

The comparison of the correlation between the CB-SIT and WAIS FS-IQ's with the correlation of the SIT and WAIS FS-IQ's resulted in a non-significant _t_ ratio of .23, _df_=45. In addition, similar analyses showed no significant differences in the correlations between CB-SIT and SIT scores with WAIS VIQ's or PIQ's. Thus, the magnitudes of the validity relationships were essentially parallel for the CB-SIT and SIT in relation to the WAIS.

Since the standard deviations for the entire sample (see Table 1) reflect a restriction in range, estimated validity coefficients were derived for the entire range of subjects (Guilford, 1965). The resulting correlations of .92 between CB-SIT and WAIS FS-IQ's and .89 between SIT and WAIS FS-IQ's demonstrated that the correlations of .54 and .52 underestimated the concurrent validity between these tests. However, the correlations of .54 and .52 probably adequately reflect the nature and magnitude of the relationship for a college population.

Due to the significant order of test administration effect, separate correlational analyses were computed for each of the three testing orders. These correlational coefficients are presented in Tables 3, 4, and 5. In Groups 1 and 2, CB-SIT and SIT scores correlated positively with WAIS FS-IQ's. However, these correlations were not significant for Group 2. The patterns of correlational relationship seem to be consistent with the CB-SIT correlating about equally as the SIT with WAIS VIQ's, PIQ's, and FS-IQ's.

TABLE 3.--Coefficients of Correlation Between WAIS, SIT, and CB-SIT for Group I (N=16)

TEST	CB-SIT	SIT	WAIS VERBAL IQ	WAIS PERF. IQ	WAIS FS-IQ
CB-SIT		.77**	.42	.39	.51*
SIT			.49	.42	.60*

* $p < .05$

** $p < .01$

TABLE 4.--Coefficients of Correlation Between WAIS, SIT, and CB-SIT for Group II (N=16)

TEST	CB-SIT	SIT	WAIS VERBAL IQ	WAIS PERF. IQ	WAIS FS-IQ
CB-SIT		.79**	.54*	-.05	.36
SIT			.45	.07	.37

* $p < .05$

** $p < .01$

TABLE 5.--Coefficients of Correlation Between WAIS, SIT, and CB-SIT for Group III (N=16)

TEST	CB-SIT	SIT	WAIS VERBAL IQ	WAIS PERF. IQ	WAIS FS-IQ
CB-SIT		.75**	.75**	.55*	.82**
SIT			.69**	.52*	.76**

[illegible]

[illegible]

DISCUSSION

The present research focused on developing and evaluating a computer-based administration and scoring program for the Slosson Intelligence Test (Slosson, 1963). Although the feasibility of such a technological approach was previously demonstrated for the administration of the WAIS, the real-time scoring as well as the administration of an intelligence test has yet to be documented. The significant feature of the approach reported in this paper is that the entire diagnostic evaluation occurs within a terminal-oriented interaction between the student and the computer system.

The feasibility of this automated approach can be seen from the relationships of CB-SIT scores and SIT scores for the same subjects. The magnitude of these correlations were in the order of $r = .75$ ($p < .01$) for the entire sample, as well as for the three experimental groups. It should be noted that these correlations are somewhat lower than reported by Orr (1969) for the automated administration of the WAIS in comparison to a traditional administration. However, Arnold, Stewart, and Rawson (1969) found depressed correlational relationships for the automated WAIS when dealing with a homogeneous sample of college students. A similar restriction in range may account for the lower relationships found in the present investigation. In addition, the inclusion of a typing skill factor within the CB-SIT may also partly account for the observed correlational differences between the CB-SIT and the SIT.

With respect to the intelligence scores, since the intertest interval was about one week on the average, some memory effects are to be expected. This expectation was confirmed in finding a significant order effect of test administration.

There was no apparent effect of order with respect to WAIS as can be noted by the equivalent mean WAIS FS-IQ's in all groups. However, the effects of order are clearly indicated in all three groups for the different administration methods for the SIT. When the CB-SIT preceded the SIT (i.e., Groups 1 and 2), a mean IQ increase of approximately 10 points was noted. In contrast, when the SIT preceded the CB-SIT (Group 3) a mean IQ decrease of approximately 7 points was found.

The observed increase in Groups 1 and 2 may partially be a function of the stimulus characteristics of the CB-SIT item presentation. Items are visually presented on the CRT, whereas, in a traditional SIT administration, item presentation is verbal in nature. Furthermore, the CB-SIT items remain on the screen until the student responds to the item. Thus, the nature of this visual presentation coupled with increased presentation time may have reduced the students' memory requirements for each item, and therefore, maximized practice effects. In turn, this memory effect might explain the non-significant correlation coefficients between the CB-SIT and the WAIS found in Groups 1 and 2. The unexpected decrease in Group 3 may be a function of both the lack of the visual memory effects as well as the observed lower scores for CB-SIT scores which might be due to the necessary typing skill for response input.

A major question concerns the adequacy of the word/phrase dictionary approach for scoring and deriving valid IQ scores for the CB-SIT. An analysis of psychologist-derived versus computer-derived IQ scores yielded a correlation of .89 for the entire sample of 48 college students. This is to be contrasted with the initial field trial results which showed a correlation of .83 between the computer-derived and psychologist-derived IQ scores. The mean CB-SIT derived IQ score was 121.2 (sd=8.56); in comparison, the mean psychologist-derived IQ score was 125.7 (sd=9.6). The 11 point difference observed in the initial field trial between the two scoring methods was reduced to 4.5 IQ points in the present study.

Thus, our scoring procedures for the CB-SIT still tend to produce somewhat lower estimates of IQ scores than the SIT. However, this latter difference is almost within the standard error of measurement for the traditional SIT (SEM=4.3). As will be seen from the validity data, this underestimation appeared to be constant per subject, and therefore, did not reduce the validity relationships of the CB-SIT in terms of the WAIS. The correlations of both tests with the WAIS did not differ significantly from each other (r=.54 for the CB-SIT and .52 for the SIT). Thus, the feasibility and validity of an automated administration and scoring program for the SIT are clearly demonstrated for this college population.

This automated intelligence testing procedure may increase the predictive efficiency of the SIT in comparison to scores obtained from traditional administration procedures. Two major considerations lead to the formulation of this rationale. First, preliminary evidence from our laboratory provides some indication that computer evaluation may be less threatening than instructor evaluation (Gallagher, 1970). The reduction of certain affective variance components which adversely effect test scores should lead to increased reliability and validity of the obtained IQ scores (Cronbach, 1960). We are actively investigating the affective nature and consequences of this man-machine interactive testing experience.

Second, the addition of multiple response indices (latency, time to complete the test, etc.) should incorporate new dimensions in the IQ derivation process. We plan to evaluate the relationships between certain temporal response measures and their relationships to the automated CB-SIT scores. Statistical prediction models will then be developed to include these relationships and evaluated with reference to college grade point average.

FOOTNOTES

1. The authors wish to express their appreciation to Richard L. Slosson for his permission to computerize the Slosson Intelligence Test. In addition, the authors would like to thank Mrs. Barbara Johnson for her editorial assistance. This research was supported in part by grants to the second author from the United States Office of Education (OEG-0-70-2671), and also to the third author from the Office of Naval Research (N00014-68-A-0494).
2. These data will be reported in a subsequent paper.

REFERENCES

Arnold, O. J., Stewart, L. T., & Rawson, H. E. Some exploratory research on the functional usage of the automated testing booth. Paper presented at the meeting of the Indiana Psychological Association, Indianapolis, April, 1969.

Cronbach, L. J. Essentials of psychological testing. (2nd ed.) New York: Harper & Row, 1960.

Elwood, D. L. Automation of psychological testing. American Psychologist, 1969, 24(3), 287-289.

Gallagher, P. D. An investigation of instructional treatments and learner characteristics in a computer-managed instruction course. Technical Report No. 12, Computer-Assisted Instruction Center, Florida State University, 1970.

Griffith, C. R., & Elwood, D. L. The automated administration of the Wechsler Adult Intelligence Scale. Paper presented at the meeting of the Indiana Psychological Association, Indianapolis, April, 1969.

Guilford, J. P. Fundamental statistics in psychology and education. (4th ed.) New York: McGraw-Hill, 1965.

International Business Machines Corporation. IBM 1500 Coursewriter II author's guide. New York: IBM, 1967.

McNemar, Q. Psychological statistics. (3rd ed.) New York: John Wiley & Sons, Inc., 1962.

Orr, T. B. A comparison of the automated method and the face-to-face method of administering the Wechsler Adult Intelligence Scale. Paper presented at the meeting of the Indiana Psychological Association, Indianapolis, April, 1969.

Roget, P. (ed.) Roget's international thesaurus. New York: Crowell, 1962.

Slosson, R. L. The Slosson Intelligence Test for Children and Adults. East Aurora: Slosson Education Publication, 1963.

Soule, R. Soule's dictionary of English synonyms. Boston: Little, Brown, & Co., 1966.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. The state-trait anxiety inventory (Test Manual). Palo Alto, Calif: Consulting Psychologists Press. In Press, 1970.

Vernon, P. E. Education and the psychology of individual differences. Harvard Educational Review, 1958, 28, 91-104.

Wechsler, D. Wechsler Adult Intelligence Scale Manual. New York: Psychological Corporation, 1955.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.